Supplementary Information

Experimentally-Determined Strengths of Favorable, Unfavorable Interactions of Amide Atoms Involved in Protein Self-Assembly in Water

Xian Cheng,[1,2] Irina A. Shkel,[2,3] Kevin O'Connor[2,] and M. Thomas Record Jr.[1,2,3]

Program in Biophysics[1] and Departments of Biochemistry[2] and Chemistry[3]
University of Wisconsin – Madison, Madison WI 53706

Supplementary Information is organized into five sections. Section I estimates two-way alpha values for atom-atom interactions of amide compounds from one-way alpha values for amide pairs differing primarily in ASA of one atom type, and compares results of this direct but approximate approach with those from the global fitting analysis. Section II provides several applications of two-way alpha values to predict or interpret effects of amides and polyamides on biopolymer processes, including effects of urea and polyvinyl pyrrolidone (PVP) on protein unfolding, and potential applications to protein self-assembly interactions. Section III provides specific examples of one- and two- way dissections of μ_{23} values for amide-amide interactions. Section IV provides consistency checks with two-way and one-way alpha values and compares analyses of the amide sp^2C and aromatic sp^2C data sets. Section V provides details of Materials and Methods.

www.pnas.org/cgi/doi/10.1073/pnas.2012481117

I) Estimates of Two-way Alpha Values for Amide Atom-Atom Interactions from Differences in One-way Alpha Values for Amide Pairs Differing Primarily in ASA of One Atom Type: A Direct Demonstration of the Validity of the Global Fitting Analysis

Two-way alpha values, reported in Table 1 from global fitting of 105 amide-amide and amide-aromatic μ_{23} values to Eq. 3, can be estimated directly from comparisons of one-way alpha values $\alpha_{3,i}$ for selected pairs of amide solutes that differ primarily in ASA of one atom type (j). For example, ethylurea and 1,1-diethylurea differ from methylurea and 1,1-dimethylurea primarily in aliphatic sp^3C ASA. The sp^3C ASA difference between the dialkylated ureas (64 $Å^2$) is about twice as large for the monoalkylated ureas (36 $Å^2$). These sp^3C ASA differences are 85-90% of the total magnitude of ASA differences for these pairs of compounds. 1,3-diethylurea (1,3-deu) differs from propionamide primarily in amide sp^2N ASA (20 $Å^2$, which is 84% of the total magnitude of ASA differences for this pair of compounds). Finally, the diamide N-acetylalanine N-methylamide (aama) differs from 1,3-deu primarily in amide sp^2O ASA (34 $Å^2$, which is 63% of the total magnitude of ASA differences for this pair of compounds).

Previously, for these same pairs of amide compounds, we analyzed differences in μ_{23} values ($\Delta\mu_{23}$) for interactions with urea and the series of alkyl ureas and found regular progressions with increasing alkylation of the urea that matched those of the corresponding one way alpha values (1). These $\Delta\mu_{23}$ values, divided by the corresponding ASA differences, provide estimates of these one-way alpha values. Here, we extend this analysis to differences in one-way alpha values $\Delta\alpha_{3,i}$ for these pairs of amide solutes differing in one type of ASA. This analysis provides semi-quantitative estimates of the corresponding two-way alpha values α_{ij}, using ASA differences for atom type j ($\Delta ASA_{j(3)}$):

$$\boldsymbol{\alpha_{ij} \cong \Delta\alpha_{3,i}/\Delta ASA_{j(3)}} \qquad \textbf{Eq. S1}$$

An example of this calculation is provided in a subsequent SI section (Eqs. S10-13) to clarify the notation.

Applying Eq. S1 to estimate two-way alpha values for the interactions of aliphatic sp^3C atoms with amide sp^2O, N and C atoms from differences in the corresponding one-way alpha values for pairs of dialkyl and monoalkyl amide compounds and their differences in sp^3C ASA yields the direct estimates in Table S7. For the dialkyl ureas, where the differences in one-way alpha values are larger and better determined, estimates of two-way alpha values for interactions of sp^3C with all four atom types agree quantitatively with those obtained from global fitting, differing by less than the propagated uncertainty in the fitted values. For the monoalkyl ureas, agreement is semiquantitative, with deviations of 25-40%.

For the amide pair 1,3-diethylurea and propionamide, differing primarily in amide sp^2N ASA, Table S7 reveals near-quantitative agreement of two-way alpha values estimated from Eq. S1 with those in Table 1 obtained from fitting the entire μ_{23} dataset to Eq. 3 for interactions of amide sp^2N with amide sp^2O and aliphatic sp^3C, while estimates for interactions with amide sp^2N and amide sp^2C agree with best-fit two-way alpha values in sign but not in magnitude. With the exception of the strongly favorable amide sp^2N - amide sp^2O interaction, atom-atom interactions involving amide sp^2N are weak, which affects the ability to determine them by this difference approach. For the diamide N-acetylalanine N-methylamide (aama) and the amide 1,3-diethylurea (1,3-deu), differing primarily in amide sp^2O ASA, estimates of two-way alpha values for interactions of sp^2O from differences in one-way alpha values (Eq. S1), listed in Table S7, agree with best-fit two-way alpha values (Table 1) within 10-30%. This level of agreement is obtained because interactions involving sp^2O are the strongest of any atom type, and hence are dominant even for this situation where sp^2O is only two-thirds of the total ASA difference between the two amide compounds.

II) Using Two-Way Alpha Values to Predict Amide, Polyamide Effects on Biopolymer Processes

A) Predicting Chemical Contributions to Interactions of the Polyamide PVP with Protein Surfaces and Effects of PVP on Protein Processes; Comparison with PEG

The water-soluble, flexibly-coiling polyamide polyvinylpyrrolidone (PVP), available in several different molecular weight ranges, has occasionally been used in place of the polyether PEG (polyethylene glycol) as a "macromolecular crowder" in studies of protein stability and interactions under conditions of high volume occupancy. Two-way alpha values from Table 1 are useful to predict the chemical interaction of PVP and its model monomer (N-ethyl pyrrolidone, NEP) with proteins and compare PVP with PEG. For PEG, where a wide range of molecular weights from monomer to oligomers and polymers is available, we previously determined the chemical interactions of end and interior groups of PEG with the different types of protein atoms (2) and separated chemical (preferential interaction) and physical (excluded volume) effects of PEG oligomers and polymers on protein (3) and nucleic acid (4) processes. Short oligomers of PVP are not commercially available to determine preferential interactions with protein atoms as done for PEG, but since NEP and PVP are amides, their chemical interactions with protein atoms and their chemical effects on protein processes can be predicted from the two-way alpha values obtained here and reported in Table 1.

For polymeric PVP with an average degree of polymerization N_3 greater than about 20 residues, the per-residue interaction with another solute (component 2), i.e. μ_{23}^{pi}/N_3, is well approximated as the interaction of an interior PVP residue, neglecting differences between end and interior residues:

$$\mu_{23}^{pi}/N_3 \cong \sum_i \alpha_{I,i}\, ASA_{i(2)} \qquad \text{Eq. S2}$$

The corresponding expression for PVP oligomers where contributions from the end groups should be treated separately is given in Eq. S3 below. In Eq. S2, $\alpha_{I,i}$ quantifies the strength of interaction of an interior PVP residue with the i-th type of atom on solute 2, with accessible

surface area $ASA_{i(2)}$. While neither these $\alpha_{I,i}$ values for PVP nor the corresponding $\alpha_{NEP,i}$ values for NEP have been determined directly, both are readily predicted using Eq. 2 from the two-way alpha values in Table 1 and ASA information for PVP interior residues and NEP. Table S8 summarizes this ASA information, and Table S9 lists predicted one-way alpha values ($\alpha_{I,i}$, $\alpha_{NEP,i}$ and the PVP end-group one-way alpha value $\alpha_{E,i}$) and compares these PVP residue one-way alpha values with those obtained previously (2) for PEG residues.

Table S9 shows a similar pattern of interactions of PVP and PEG interior residues with the most common types of protein atoms. Chemical interactions of both PVP and PEG residues with protein aliphatic sp^3C, amide sp^2N and amide sp^2C are favorable, while interactions with protein amide sp^2O are unfavorable. Interactions of PVP interior residues with aliphatic sp^3C and amide sp^2O are about 2-fold stronger than for PEG, while PVP interactions with protein amide sp^2C and amide sp^2 N are about 1.6- and 1.2-fold stronger than for PEG.

Table S10 predicts that PVP monomer (NEP) destabilizes globular proteins, in agreement with its observed destabilization of CI2 (5). NEP is predicted to stabilize α-helices; this difference results from the very different composition of the ASA exposed in unfolding α-helices (predominantly amide, weighted toward amide sp^2O) as compared to unfolding globular proteins (predominantly sp^3C). By contrast, PEG monomer (ethylene glycol) and dimer (diethylene glycol) are predicted and observed to stabilize globular proteins, while tri- and tetraethylene glycol are destabilizing (2). Polymeric PVP is observed to stabilize CI2 (6, 7), which we conclude is because the predicted stabilizing excluded volume effect (3) counterbalances the predicted small destabilizing chemical effect (see section C) below and Table S10).

Hence PVP exerts chemical effects which differ only modestly from those of PEG. Since both PVP and PEG are flexible polymers with similar persistence lengths, their excluded volume

contributions to μ_{23} are also expected to be similar. Since PEG is available at high purity over a much wider range of chain lengths than PVP, it is the better choice for these studies.

Eq. S2 is the high-polymer limiting form of the general expression for chemical (preferential interaction) contributions to μ_{23} values for interactions of PVP oligomers or polymers of any number of residues (N_3) with the different hybridization states of O, N and C atoms of other solutes or proteins. Since any PVP has two end residues and N_3-2 interior residues, the interaction of the average PVP residue with a solute 2 is therefore (3)

$$\frac{\mu_{23}^{pi}}{N_3} = \left(1 - \frac{2}{N_3}\right) \sum_i \alpha_{I,i}\, ASA_{i(2)} \; + \left(\frac{2}{N_3}\right) \sum_i \alpha_{E,i}\, ASA_{i(2)} \qquad \textbf{Eq. S3}$$

For N_3 > 20, Eq. S3 reduces to Eq. S2. In Eq. S3, $\alpha_{E,i}$ and $\alpha_{I,i}$ are one-way alpha values that quantify the intrinsic strength of interactions of PVP end (E) and interior (I) residues with the i-th type of atom on another solute or protein. For PEG, where the end residues (as defined) are half the size of interior residues, we combined them ($\alpha_{2E,i}$) but for PVP it is more appropriate to treat each end residue separately. In Eq. S2 for μ_{23} for high molecular weight PVP (N_3 >> 1)), no distinction is made between end and interior residues.

By analogy with Eq. 2, each one-way $\alpha_{E,i}$ and $\alpha_{I,i}$ in Eq. S3 is itself a sum of contributions of interaction of the i-th type of protein atom with the j-th type of PVP atom (see Eq.3).

$$\alpha_{E,i} = \sum_j \alpha_{ij}\, ASA_{j(PVP,E)} \qquad \text{and} \qquad \alpha_{I,i} = \sum_j \alpha_{ij}\, ASA_{j(PVP,I)} \qquad \textbf{Eq. S4}$$

In Eq. S4, each two-way alpha value (α_{ij}) quantifies the interaction of 1 Å^2 atom type i of solute 2 with 1 Å^2 of PVP atom type j, $ASA_{j(PVP,E)}$ and $ASA_{j(PVP,I)}$ are areas (in Å^2) of atom type j on the end and interior residues of PVP. One-way alpha values for NEP and for PVP end and interior residues, calculated from two-way alpha values as in Eqs. 2 and S4, are listed in Table S9 and compared with the corresponding quantities for PEG.

For interactions of larger PVP oligomers and polymers with large solutes, an excluded volume term also contributes to μ_{23} and the chemical term in Eq. S3 may be reduced by a shielding term χ (3). As previously (3) , we interpret the interaction of PVP (component 3) with a protein (component 2) as the sum of preferential interaction (abbreviated pi) and excluded volume (ev) contributions,

$$\mu_{23} = \chi\, \mu_{23}^{pi} + \mu_{23}^{ev} \qquad \textbf{Eq. S5}$$

In Eq. S5, the quantity χ is the fractional accessibility of the average residue of PVP. For a PVP oligomer one expects $\chi \approx 1$, but for polymeric PVP one expects $\chi \ll 1$. (3)

B) Predicting m-Values for *Urea and other Amide Solutes.* Another significant application of two-way alpha values for amide compound atom-atom interactions is to predict or interpret effects (*m*-values) of urea or any other amide solute on protein processes (8, 9) in terms of ASA information for the amide solute and ΔASA information for each type of unified protein atom using Eqs. 2 and 3. Recently, we analyzed urea *m*-values for unfolding of globular proteins using urea one-way alpha values for amide sp^2O, amide sp^2N, aliphatic sp^3C and aromatic sp^2C and ΔASA information assuming an extended chain model of the unfolded state. Generally good agreement is obtained between experimental *m*-values and *m*-values predicted either using these four major protein atom types or using all seven protein atom types (8). Fig. S3 shows that use of two-way alpha values from Table 1 yields predicted *m*-values which agree well with the two one-way predictions and with experimental *m*-values.

C) Applications of two-way alpha values to protein self-assembly interactions

Potential new directions of research using two-way alpha values include predicting or interpreting χ parameters of Flory-Huggins theory (10-12) in applications to aqueous polymer solutions. In this theory, χ quantifies the strength of segment-water interactions relative to

segment-segment and water-water interactions. Extensions of this theory with “stickers and spacers” provide more realistic analyses of interactions of segments of flexible chain models of biopolymers (12, 13). It seem likely that two-way alpha values for amide atoms can be used to predict or interpret χ and noncoulombic “sticker" and “spacer” interaction parameters in analyses of the different behaviors of low-complexity polypeptides and unfolded proteins in chain expansion-collapse and aggregation (12, 14-16). Use of two-way alpha values would allow the sticker and spacer treatment to be extended to include a third type of region with net-unfavorable interactions (positive alpha values). Expansion of the set of two-way alpha values to include interactions of protein and nucleic acid unified atoms will allow their use in coarse-grained simulations and other analyses of interactions in liquid droplets formed by RNA and RNA binding proteins (12, 13, 16-18).

III) Specific Examples of One- and Two- Way Dissections of μ_{23} Values for Amide-Amide Interactions

In this section we provide specific expressions applying Eqs. 1-3 and S1 to the amides studied here.

A) $$\mu_{23} = \sum_i \alpha_{3,i} ASA_{i(2)}$$ **Eq. 1**

Each contribution in this sum is composed of an intrinsic interaction strength (one-way alpha value) for the interaction of solute *3* with a unit area (1 Å^2) of accessible surface of one type of unified atom of the biopolymer or other solute 2. For the interactions of amide compounds investigated here, these atom types are amide sp^{2}O, sp^{2}N and sp^{2}C, and aliphatic sp^{3}C. Taking as a specific example the interaction of N-acetylalanine N-methylamide (aama, component 3) with the various amide atoms of propionamide (ppa, component 2), for which μ_{23}= $\mu_{ppa,aama}$ = -43 $\pm$ 4 cal mol^{-1} molal^{-1} (Table S1):

$$\mu_{ppa,aama} = -43 \pm 4 = \alpha_{aama,sp^2O} ASA_{sp^2O(ppa)} + \alpha_{aama,sp^2N} ASA_{sp^2N(ppa)} + \alpha_{aama,sp^2C} ASA_{sp^2C(ppa)} + \alpha_{aama,sp^3C} ASA_{sp^3C(ppa)}$$

$$= \mathbf{36.8}\,\boldsymbol{\alpha}_{\boldsymbol{aama,sp^2O}} + \mathbf{61.6}\,\boldsymbol{\alpha}_{\boldsymbol{aama,sp^2N}} + \mathbf{4.3}\,\boldsymbol{\alpha}_{\boldsymbol{aama,sp^2C}} + \mathbf{124}\,\boldsymbol{\alpha}_{\boldsymbol{aama,sp^3C}} \quad \textbf{Eq. S6}$$

where the propionamide ASA values (1) in Eq. S6 are in Å^2 and the units of the one-way alpha values are cal mol^{-1} molal^{-1} Å^{-2}.

Ten other equations like Eq. S6 are written to interpret experimental $\mu_{2,aama}$ values for the interactions of aama with formamide, N-methylformamide, malonamide, urea, methylurea, and the remainder of the set of amide compounds (component 2) investigated. Solving these eleven equations in four unknowns (α_{aama,sp^2O}, α_{aama,sp^2N}, α_{aama,sp^2N}, α_{aama,sp^3C}) determines best-fit values for the above four one-way aama alpha values (Table S2). Comparison of predicted and observed $\mu_{2,aama}$ values for the full set of eleven aama-amide compound interactions (Fig. 2B, Table S3) tests the hypotheses of additivity and proportionality of contributions to ASA which are the basis of Eqs 1 and S6. Analogous sets of eleven equations are formulated and solved to obtain sets of four one- way alpha values quantifying the interactions of each other amide compound (formamide, N-methylformamide, malonamide, propionamide) with a unit area of amide sp^{2}O, sp^{2}N, sp^{2}C and aliphatic sp^{3}C atoms (Table S2).

Since μ_{23}= μ_{32} therefore

$$\boldsymbol{\mu_{aama,ppa}} = \mathbf{-43 \pm 4} = \boldsymbol{\alpha_{ppa,sp^2O}ASA_{sp^2O(aama)}} + \boldsymbol{\alpha_{ppa,sp^2N}ASA_{sp^2N(aama)}} + \boldsymbol{\alpha_{ppa,sp^2C}ASA_{sp^2C(aama)}} + \boldsymbol{\alpha_{ppa,sp^3C}ASA_{sp^3C(aama)}}$$

$$= \mathbf{62.5}\boldsymbol{\alpha_{ppa,sp^2O}} + \mathbf{21.1}\,\boldsymbol{\alpha_{ppa,sp^2N}} + \mathbf{4.3}\,\boldsymbol{\alpha_{ppa,sp^2C}} + \mathbf{258}\,\boldsymbol{\alpha_{ppa,sp^3C}} \quad \textbf{Eq. S7}$$

Ten other equations like Eq. S7 are written to interpret experimental $\mu_{2,ppa}$ values for the interactions of propionamide with formamide, N-methylformamide, malonamide, urea, methylurea, and the remainder of the set of amide compounds (component 2) investigated. Solving these eleven equations in four unknowns (α_{ppa,sp^2O}, α_{ppa,sp^2N}, α_{ppa,sp^2N}, α_{ppa,sp^3C}) determines best-fit values for the above four one-way propionamide alpha values (Table S2).

B) $$\boldsymbol{\alpha_{3,i}} = \sum_j \boldsymbol{\alpha_{ij}ASA_{j(3)}} \quad \textbf{Eq. 2}$$

Continuing with the above example for the ppa-aama interaction,

$$\alpha_{aama,sp^2O} = \mathbf{2.88 \pm 0.21\ cal\ mol^{-1} molal^{-1} Å^{-2}} = \alpha_{sp2O,sp2O} ASA_{sp2O\,(aama)} + \alpha_{sp2O,sp2N} ASA_{sp2N(aama)} + \alpha_{sp2O,sp2C} ASA_{sp2C(aama)} + \alpha_{sp2O,sp3C} ASA_{sp3C(aama)}$$

$$= 62.5\alpha_{sp2O,sp2O} + 21.1\alpha_{sp2O,sp2N} + 4.3\alpha_{sp2O,sp2C} + 258\alpha_{sp2O,sp3C} \quad \textbf{Eq. S8}$$

$$\alpha_{ppa,sp^2O} = \mathbf{1.04 \pm 0.09\ cal\ mol^{-1} molal^{-1} Å^{-2}} = \alpha_{sp2O,sp2O} ASA_{sp2O\,(ppa)} + \alpha_{sp2O,sp2N} ASA_{sp2N(ppa)} + \alpha_{sp2O,sp2C} ASA_{sp2C(ppa)} + \alpha_{sp2O,sp3C} ASA_{sp3C(ppa)} =$$

$$36.8\,\alpha_{sp2O,sp2O} + 61.6\,\alpha_{sp2O,sp2N} + 4.3\,\alpha_{sp2O,sp2C} + 124\,\alpha_{sp2O,sp3C} \quad \textbf{Eq. S9}$$

where the units of the two-way alpha values are cal mol^{-1} molal^{-1} Å^{-4}. Nine other equations like Eqs. S8-9 are written to interpret experimental $\alpha_{3,sp2O}$ values for the interactions of formamide, N-methylformamide, malonamide, urea, methylurea, and the remainder of the set of amide compounds (component 2) with sp^{2}O atoms. Solving these eleven equations in four unknowns ($\alpha_{sp2O,sp2O}, \alpha_{sp2O,sp2N}, \alpha_{sp2O,sp2C}, \alpha_{sp2O,sp3C}$) determines best-fit values for these four two-way alpha values (Table 1). Comparison of predicted and observed μ_{23} values for the full set of eleven aama-amide compound interactions (Fig. 2B, Table S3) tests the hypotheses of additivity and proportionality of contributions to ASA which are the basis of Eqs 2 and S8-9. An analogous set of eleven equations is formulated and solved to obtain a set of four two- way alpha values quantifying the interactions of a unit area of each other type of unified atom (amide sp^{2}N, sp^{2}C and aliphatic sp^{3}C) with a unit area of amide sp^{2}O, sp^{2}N, sp^{2}C and aliphatic sp^{3}C atoms (Table 1).

C) $$\alpha_{ij} = \Delta\alpha_{3,i}/\Delta ASA_{j(3)} \quad \textbf{Eq. S1}$$

Here we illustrate the application of Eq. S1 to one of the four pairs of amides (methylurea and ethylurea) analyzed in the text. These amides differ primarily in amount of sp^{3}C ASA. (See text for a discussion of all four amide pairs, based on the numerical analysis in Table S7.)

For methylurea (mu), the one-way alpha value for the interaction with 1 Å^2 of amide sp^{2}O surface (0.78 cal mol^{-1} molal^{-1} Å^{-2}; (1)) is interpreted by Eq. 2 as the sum of ASA-weighted contributions from interactions of the methyl sp^{3}C and amide sp^{2}O, sp^{2}N and sp^{2}C atoms of methylurea (*mu*) with amide sp^{2}O atoms of other compounds:

$$\alpha_{mu,sp2O} = \alpha_{sp3C,sp2O}ASA_{sp3C(mu)} + \alpha_{sp2O,sp2O}ASA_{sp2O(mu)} + \alpha_{sp2N,sp2O}ASA_{sp2N(mu)} + \alpha_{sp2C,sp2O}ASA_{sp2C(mu)} \quad \textbf{Eq. S10}$$

The corresponding equation for ethylurea (eu) is

$$\alpha_{eu,sp2O} = \alpha_{sp3C,sp2O}ASA_{sp3C(eu)} + \alpha_{sp2O,sp2O}ASA_{sp2O(eu)} + \alpha_{sp2N,sp2O}ASA_{sp2N(eu)} + \alpha_{sp2C,sp2O}ASA_{sp2C(eu)} \quad \textbf{Eq. S11}$$

Subtracting Eq. S11 from S10 yields a specific example of $\Delta\alpha_{3,i}$:

$$\alpha_{eu,sp2O} - \alpha_{mu,sp2O} = \alpha_{sp3C,sp2O}\Delta ASA_{sp3C} + \alpha_{sp2O,sp2O}\Delta ASA_{sp2O} + \alpha_{sp2N,sp2O}\Delta ASA_{sp2N} + \alpha_{sp2C,sp2O}\Delta ASA_{sp2C} \quad \textbf{Eq. S12}$$

Because 87% of the ASA difference between ethylurea and methylurea is from sp^{3}C, to a good approximation $\alpha_{eu,sp2O} - \alpha_{mu,sp2O} \approx \alpha_{sp3C,sp2O}\Delta ASA_{sp3C}$ and

$$\alpha_{sp3C,sp2O} \approx (\alpha_{eu,sp2O} - \alpha_{mu,sp2O})/\Delta ASA_{sp3C} \quad \textbf{Eq. S13}$$

which is a specific example of Eq. S1. Table S7 summarizes the results of this and three other difference analyses to estimate two-way alpha values, and compares these estimates with those in Table 1, obtained from global fitting. For the case of $\alpha_{sp3C,sp2O}$ analyzed above, the estimate from Eq. S13 is within 30% of the Table 1 value, as shown in Table S7.

D) $$\mu_{23} = \sum_i \sum_j \alpha_{ij} ASA_i ASA_j \quad \textbf{Eq. 4}$$

As in sections A-C above, indices i and j refer to the four types of unified atom present in the amide compounds investigated (amide sp^{2}O, sp^{2}N, sp^{2}C; aliphatic sp^{3}C). Hence, for each of the one hundred and five pairs of amide compounds investigated:

$$\mu_{23} = \alpha_{sp^2O,sp^2C}\left(ASA_{sp^2O(2)}ASA_{sp^2C(3)} + ASA_{sp^2C(2)}ASA_{sp^2O(3)}\right) +$$

$$\alpha_{sp^2O,sp^2N}\left(ASA_{sp^2O(2)}ASA_{sp^2N(3)} + ASA_{sp^2N(2)}ASA_{sp^2O(3)}\right) + \alpha_{sp^2O,sp^3C}\left(ASA_{sp^2O(2)}ASA_{sp^3C(3)} +$$

$$ASA_{sp^3C(2)}ASA_{sp^2O(3)}\right) + \alpha_{sp^2O,sp^2O}ASA_{sp^2O(2)}ASA_{sp^2O(3)} + \alpha_{sp^2N,sp^3C}\left(ASA_{sp^2N(2)}ASA_{sp^3C(3)} +$$

$$ASA_{sp^3C(2)}ASA_{sp^2N(3)}\right) + \alpha_{sp^2N,sp^2C}\left(ASA_{sp^2N(2)}ASA_{sp^2C(3)} + ASA_{sp^2C(2)}ASA_{sp^2N(3)}\right) +$$

$$\alpha_{sp^2N,sp^2N}ASA_{sp^2N(2)}ASA_{sp^2N(3)} + \alpha_{sp^2C,sp^2C}ASA_{sp^2C(2)}ASA_{sp^2C(3)} +$$

$$\alpha_{sp^2C,sp^3C}\left(ASA\ _{sp^2C(2)}ASA\ _{sp^3C(3)} + ASA\ _{sp^3C(2)}ASA\ _{sp^2C(3)}\right) +$$

$$\alpha_{sp^3C,sp^3C}ASA_{sp^3C(2)}ASA_{sp^3C(3)} \qquad \textbf{Eq. S14}$$

Here, as in sections A-C above, $ASA_{i(3)}$ is the ASA of group i on solute 3 and $ASA_{j(2)}$ is the ASA of group j on solute 2.

IV) Consistency Checks with Two-way and One-Way Alpha Values

A) Predicting One-Way Alpha Values for Interactions of Amide Solutes with the Types of Unified Atoms of Amide Compounds

As described in the main text, two-way alpha values can be used to predict one way alpha values quantifying how any amide solute interacts with amide sp^2O, sp^2N, sp^2C and aliphatic sp^3C unified atoms on any amide or polyamide molecule or surface (e.g. the surface exposed in protein unfolding). As an example, one-way alpha values for interactions of all twelve amide solutes investigated with unit areas of the four amide unified atoms may be predicted from two-way alpha values (Table 1) and ASA information (1) using Eqs. 2 (see Eqs. S8-9 for examples), and compared with observed one-way alpha values determined from μ_{23} values using Eq. 1 (see Eqs. S6-7 for examples) and ASA information. Results of these two methods to obtain one-way alpha values are shown in Table S5. Agreement within the combined 1 SD uncertainties is observed for 83% of these solute-atom interactions, and all but the interaction of N-methylformamide with sp^2O agree within 2 SD.

B) Comparison of Two-way Alpha Values for Atom-Atom Interactions of Amides from Different Treatments of sp^2C

One-way alpha values for interactions of urea and alkyl ureas with amide and aromatic sp^2C were found to be similar (1). Two-way alpha values listed in Table 1 were determined by analysis of 105 μ_{23} values for amide interactions (85 amide compound-amide compound, 20 amide compound-aromatic compound) using Eq. 3, to obtain a combined two-way alpha value for amide and aromatic sp^2C. To justify this analysis, here we extend it by global fitting all μ_{23} values (105 total, including 20 for amide- aromatic hydrocarbon interactions) to Eq. S15 below which includes one global weighting factor (w_{C_a}) quantifying the relative strength of interactions of amide sp^2C as compared to aromatic sp^2C. Clearly this is an oversimplification, since in principle a different weighting factor might be needed for interactions of sp^2C with each other type of atom, but it provides a test of whether such corrections are significant. The revised version of Eq. 3 for μ_{23} is

$$\mu_{23} = \sum_i^{i\neq sp^2C} \sum_j^{j\neq sp^2C} \alpha_{ij}\, ASA_i\, ASA_j + \sum_j^{j\neq sp^2C} \alpha_{C_{am,ar},j}\, ASA_j \left(ASA_{C_{ar}} + w_{C_{am}} ASA_{C_{am}}\right) +$$

$$\alpha_{C_{am,ar},C_{am,ar}} w_{C_{am}} ASA_{C_{am}} \left(ASA_{C_{ar}} + w_{C_{am}} ASA_{C_{am}}\right) \qquad \textbf{Eq. S15}$$

In Eq. S15, the subscript $C_{am,ar}$ stands for combined amide and aromatic sp^2C, C_{am} stands for amide sp^2C and C_{ar} is aromatic sp^2C.

Two-way alpha values summarized in Table 1 were obtained for the unweighted case ($w_{C_{am}}$= 1). In Table S6 these values are compared with those obtained from a global analysis using Eq. S15 and floating $w_{C_{am}}$. Two-way alpha values obtained from this analysis are the same as in Table 1 within the uncertainty, although the percent difference in the interaction of sp^2O with sp^2O is about 80%. In this fit, the weighting coefficient $w_{C_{am}}$= 0.82, indicating that on average the intrinsic strengths of interactions of amide sp^2C with the different atom types are about 82% as large as for aromatic sp^2C.

Table S6 also compares two-way alpha values obtained from analyses of subsets of μ_{23} values with those in Table 1 and from the fit with $w_{C_{am}}$ weighting. Fitting only the 64 amide-amide μ_{23} values obtained for amides with minimal amounts of amide sp^2C to the variant of Eq. 3 with only six terms for interactions involving only sp^2O, sp^2N and sp^3C yields six two-way alpha values which agree within the combined uncertainty with those of Table 1. Fitting only the 20 amide-aromatic μ_{23} values to another variant of Eq. 3 yields two-way alpha values for interactions of sp^2C with sp^2O, sp^2N, sp^2C and sp^3C. Two-way alpha values for sp^2C-sp^2C and sp^2C-sp^3C agree with those in Table 1, while those for sp^2C-sp^2O, and sp^2C-sp^2N are both 20-30% larger in magnitude than their counterparts in Table 1, consistent with the finding of a weighting coefficient $w_{C_{am}}$ = 0.82 for interactions involving amide sp^2C.

C) Effect of Size of Dataset on Two-way Alpha Values

Table S6, discussed above, compared the separate determinations of four two-way alpha values from μ_{23} values for 20 amide-aromatic interactions and of six two-way alpha values for 85 amide-amide interactions with the ten two-way alpha values obtained from global analysis of the set of 105 μ_{23} values for amide-aromatic and amide-amide interactions, treating amide sp^2C as the same as or differently from aromatic sp^2C. Two-way alpha values obtained from these various approaches agree in most cases within the combined uncertainty. In Table S11 the effect of other reductions in the size of the μ_{23} data set are examined. This table shows there is little effect on two-way alpha values of removing all of the 14 to 26 μ_{23} values that quantify interactions of the more polar (urea, malonamide) and/or nonpolar (,3-diethylurea, aama) amides. The insensitivity of the two-way alpha values to these reductions in the set of μ_{23} values analyzed shows that even these subsets are large enough and diverse enough to determine all ten two-way alpha values. .

V) Materials and Methods

Chemicals

Formamide (>99.5%), N-methylformamide (>99%), and malonamide (>97%) were obtained from Sigma. Propionamide (>98%) was from Alfa Aesar and N-acetylalanine N-methylamide (aama, >99%) was from Bachem. All these amides were obtained in anhydrous form and used without further purification. All were dissolved in deionized water obtained from a Barnstead E-pure system (Thermo-Fischer Scientific).

Structures of Amide Compounds and ASA Calculations

Molecular structures of NEP (N-ethyl pyrrolidone) and of short oligomers of PVP (polyvinyl pyrrolidone) used for calculations of water-accessible surface areas (ASA) were predicted from the NIH Cactus website ((19); https://cactus.nci.nih.gov/translate/) as described previously (1, 20). Molecular structures of all other amide compounds investigated were obtained previously (1). In all cases, a unified atom model was used in which hydrogens are treated as part of the C or N atom to which they are bonded. ASA information for NEP and PVP oligomers was calculated using the program Surface Racer (20) with the Richards set of van der Waals radii (21) and a 1.4 Å probe radius for water. As previously (1), ASA values were obtained for four coarse-grained atom types: amide sp^2O, sp^2N, sp^2C and aliphatic sp^3C. Alternative sources of structural information (PubChem (22) and Biological Magnetic Resonance Bank (BMRB;(23)) and alternative ASA programs (VMD (24) and GetArea (25)) were compared previously with Cactus and Surface Racer (1, 26), and no significant differences were found.

Determination of ASA of End and interior Residues of PVP from Molecular Models of Short Oligomers

Water accessible surface areas (ASA) of the four types of unified atom (amide sp^2O, sp^2N, sp^2C; aliphatic sp^3C) of NEP and short PVP oligomers (number of residues $N_3 \leq 5$) were calculated using Surfracer (20). Results are given in Table S8. To determine ASA contributions

from the two end residues ($ASA_{2E,i}$) and the N_3-2 interior residues ($ASA_{I,i}$) of a PVP oligomer, ASA values for each type of atom (i) were fitted to Eq. S16:

$$ASA_{N_{3,},i} = ASA_{2E,i} + ASA_{I,i}\,(N_3 - 2), \qquad i = sp^3C,\ sp^2C,\ sp^2O,\ sp^2N \qquad \textbf{Eq. S16}$$

Values of $ASA_{2E,i}$ and $ASA_{I,i}$ obtained from these fits are reported in Table S8.

Vapor Pressure Osmometry (VPO)

VPO is used to quantify thermodynamic interactions of small solutes which are soluble and nonvolatile in water by measuring osmolality differences $\Delta Osm(m_2, m_3)$ between three component (water, solute 2, and solute 3) and two component (water and solute 2, water and solute 3) solutions. Details of the osmolality analysis were described previously (1).

$$\Delta Osm(m_2, m_3) = Osm\,(m_2, m_3) - \big(Osm\,(m_2) + Osm\,(m_3)\big)$$

$$\Delta Osm(m_2, m_3) = \frac{\mu_{23}}{RT}\, m_2 m_3 \qquad \textbf{Eq. S17}$$

Preferential interactions (μ_{23} values) of a series of urea and alkyl ureas with one another and with other amide compounds were determined previously by VPO using Eq. S17 (1). Here μ_{23} values quantifying pairwise interactions in aqueous solution between five additional amide compounds (formamide, N-methylformamide, propionamide, malonamide and aama) are determined.

For experiments involving all amides except aama, VPO experiments (Fig. 1) the molality of one amide (component 2, molality m_2) was held constant at 0.35 molal in one series and 0.6 molal in another, and the molality of the other amide (component 3, molality m_3) was varied from 0 to 0.95 molal. For experiments with aama, the aama concentration was held constant at 0.2 and 0.4 molal and the concentration of the other amide was varied from 0 to 0.95 molal, or the aama concentration was varied from 0 to 0.55 molal while the concentration of the other amide was held constant at 0.35 and 0.6 molal.

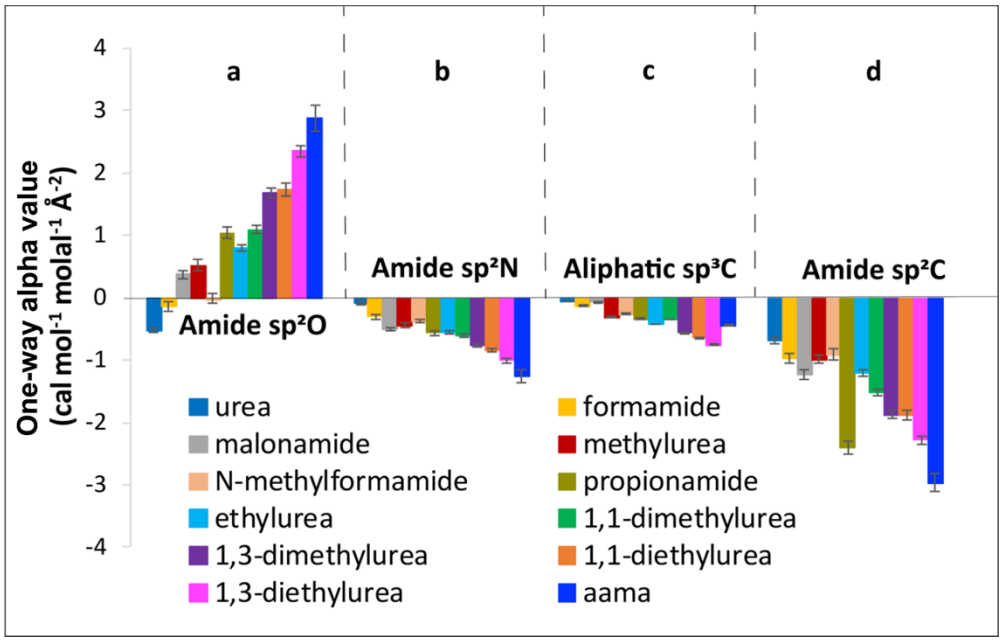

Figure S1. Comparison of One Way Alpha Values for Formamide, N-methylformamide, Malonamide, Propionamide and N-acetylalanine N-methylamide (aama) with Other Amides

Amide compounds are listed arbitrarily in order of increasing aliphatic sp^3C + amide sp^2C ASA. Bar graphs compare interaction potentials (one-way alpha values; Table S2)) quantifying interactions of each amide compound with a unit area of a) amide sp^2O, b) amide sp^2N, c) amide-context sp^3C, and d) amide/aromatic sp^2C at 23 °C. Favorable interactions have negative one-way alpha values while unfavorable interactions have positive one-way alpha values. (aama: N-acetylalanine N-methylamide). One-way alpha values for urea and alkyl ureas were reported previously (1).

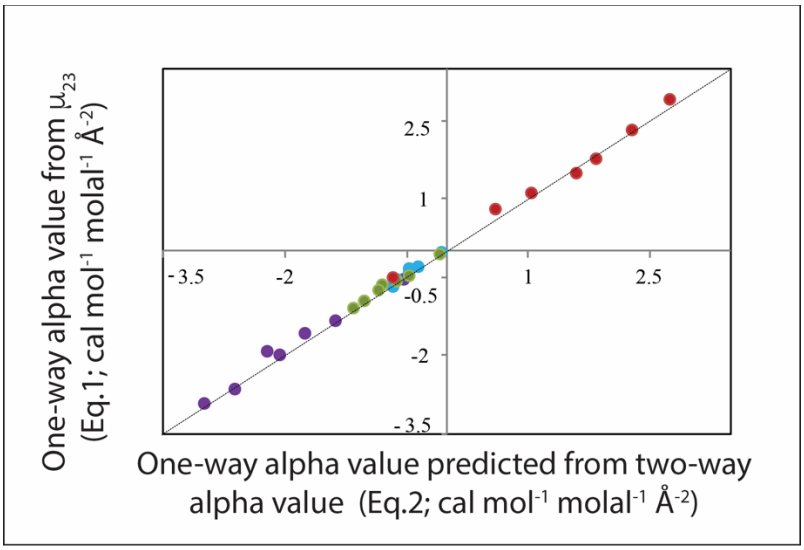

Figure S2. Comparison of Predicted and Observed One-way Alpha Values (cal mol^{-1} m^{-1} Å^{-2}) for Interactions of Urea and Alkyl ureas with Amide and Aromatic Unified Atoms (amide sp^{2}O, amide sp^{2}N, amide sp^{3}C and combined aromatic sp^{2}C and amide sp^{2}C) at 23 ^{o}C. One-way alpha values were determined by analysis of sets of μ_{23} values using Eq. 1 (1). Predictions of one-way alpha values from Eq. 2 use two-way alpha values from Table 1 and ASA information (1). The line represents equality of predicted and observed values.

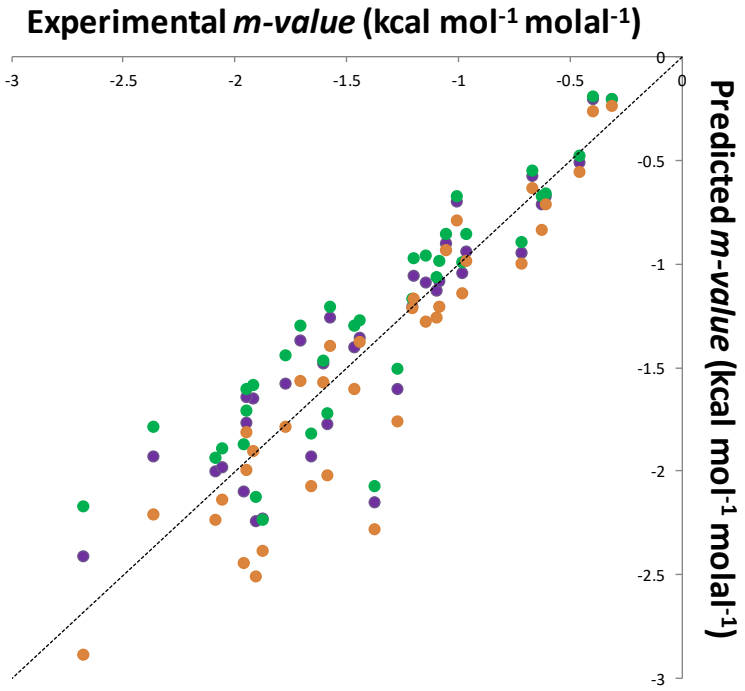

Figure S3. Comparison of Predicted and Observed Urea *m*-values of Unfolding Globular Proteins. Purple: Predicted urea *m*-values from one-way alpha values for interactions of urea with seven protein unified atoms or functional groups (8) and changes in protein ASA in unfolding (ΔASA values) (8) . Green: Predicted urea *m*-values from one-way alpha values for interactions of urea with four unified atoms of amide compounds (1) and ΔASA values (8). Yellow: Predicted urea *m*-values from two-way alpha values for atoms of amide compounds (Table 1) using Eq.3 and ΔASA values (8). Amide sp^2C represents less than 1% of the ΔASA of unfolding and was not accounted for in these comparisons.

Supplemental Tables

Table S1. Interactions of Amides with Relatively Large sp^2C and/or sp^2O Surface Area

Amide Compounds		Observed μ_{23} value[a]
propionamide	malonamide	-37.3 ± 2.2
propionamide	N-acetylalanine N-methylamide	-43.2 ± 3.9
propionamide	formamide	-76.8 ± 2.6
propionamide	N-methylformamide	-102.3 ± 1.9
malonamide	N-acetylalanine N-methylamide	-8.6 ± 2.2
malonamide	formamide	-54.1 ± 2.4
malonamide	N-methylformamide	-64.5 ± 3.0
N-acetylalanine N-methylamide	formamide	-53.0 ± 3.1
N-acetylalanine N-methylamide	N-methylformamide	-80.3 ± 3.2
formamide	N-methylformamide	-64.0 ± 2.2

[a] Values of $\mu_{23} = \mu_{32}$ determined by VPO at 23°C. Units of μ_{23} are cal mol^{-1} $molal^{-1}$.

Table S2. One-way Alpha Values for Interactions of Formamide, Malonamide, N-methylformamide, Propionamide and aama[a] With Amide sp^2O, sp^2N, sp^2C and Aliphatic sp^3C Atoms

	Amide sp^2O	Amide sp^2N	Aliphatic sp^3C	Amide sp^2C
formamide	-0.13 ± 0.08	-0.33 ± 0.03	-0.14 ± 0.01	-0.92 ± 0.08
malonamide	0.37 ± 0.07	-0.5 ± 0.02	-0.07 ± 0.01	-1.23 ± 0.07
N-methylformamide	-0.02 ± 0.09	-0.41 ± 0.03	-0.27 ± 0.01	-0.81 ± 0.09
propionamide	1.04 ± 0.09	-0.56 ± 0.04	-0.33 ± 0.01	-2.39 ± 0.1
aama[a]	2.88 ± 0.21	-1.26 ± 0.09	-0.45 ± 0.01	-2.97 ± 0.14

[a] aama: N-acetylalanine N-methylamide.

[b] One-way alpha values are obtained by fitting 11 experimental μ_{23} values for each compound listed (see Tables S1, S3) to Eq.1. Uncertainties in alpha values are calculated as previously described(2).

Table S3. Comparison of Observed μ_{23} Values (cal mol^{-1} molal^{-1}; $\mu_{23} = \mu_{32}$) for Amide Interactions at 23°C with Predictions of μ_{23} and μ_{32} from One-way Alpha Values[a]

	Malonamide (Solute 3)				**Propionamide** (Solute 3)		
Solute 2	Observed μ_{23}[a]	Predicted μ_{32} or μ_{23} from one-way alpha values of Solute 2[b]	Solute 3[c]	Solute 2	Observed μ_{23}[a]	Predicted μ_{32} or μ_{23} from one-way alpha values of Solute 2[b]	Solute 3[c]
urea	-54.9 ± 2.5	-54.4 ± 3.8	**-55.6 ± 7**	urea	-36.8 ± 2	-35.8 ± 2	**-40.7 ± 10**
mu	-43.8 ± 1.4	-43.8 ± 6.8	**-43.3 ± 6**	mu	-62.3 ± 1.7	-51.3 ± 3.9	**-54.6 ± 8.2**
eu	-48.5 ± 2.4	-46.6 ± 4.9	**-43.1 ± 6.2**	eu	-67.6 ± 1.7	-62.8 ± 2.8	**-63.6 ± 8.3**
1,1-dmu	-30.6 ± 1.8	-33.2 ± 5.3	**-30.7 ± 5.7**	1,1-dmu	-53.5 ± 2.3	-47.3 ± 3.1	**-55.3 ± 7.5**
1,3-dmu	-27.3 ± 2.4	-27.4 ± 6.9	**-31.1 ± 5**	1,3-dmu	-68.5 ± 1.9	-63 ± 4	**-68.4 ± 6.5**
1,1-deu	-34.6 ± 1.3	-37.1 ± 7.5	**-30.7 ± 5.7**	1,1-deu	-78.2 ± 1.5	-75.6 ± 4.4	**-69.8 ± 7.4**
1,3-deu	-21.7 ± 1.2	-25.8 ± 7.6	**-30.7 ± 5.3**	1,3-deu	-72.7 ± 2.5	-79.7 ± 4.5	**-86.5 ± 6.7**
ppa	**-37.3 ± 2.2**	**-37.6 ± 11.9**	**-30.8 ± 5.4**	**mad**	**-37.3 ± 2.2**	**-30.8 ± 5.4**	**-37.6 ± 11.9**
aama	**-8.6 ± 2.2**	**-12.9 ± 27.3**	**-10.4 ± 7.3**	**aama**	**-43.2 ± 3.9**	**-39.8 ± 15.7**	**-43.2 ± 9.3**
fad	**-54.1 ± 2.4**	**-63.9 ± 27.3**	**-65.3 ± 8.2**	**fad**	**-76.8 ± 2.6**	**-46.2 ± 15.7**	**-82.6 ± 11.4**
mfad	**-64.5 ± 3**	**-67.3 ± 27.3**	**-52.9 ± 7.2**	**mfad**	**-102.3 ± 1.9**	**-78.4 ± 18.9**	**-96.3 ± 9.7**
	Formamide (Solute 3)				**Methylformamide** (Solute 3)		
Solute 2	Observed μ_{23}[a]	Predicted μ_{32} or μ_{23} from one-way alpha values of Solute 2[b]	Solute 3[c]	Solute 2	Observed μ_{23}[a]	Predicted μ_{32} or μ_{23} from one-way alpha values of Solute 2[b]	Solute 3[c]
urea	-62.9 ± 2.2	-60.6 ± 3.5	**-56 ± 23.4**	urea	-56.1 ± 2.9	-57 ± 2.9	**-60.3 ± 23.4**
mu	-48.8 ± 3	-44.2 ± 5.6	**-52.1 ± 18.3**	mu	-53.7 ± 1.7	-57.2 ± 4.6	**-65.7 ± 18.3**
eu	-39.2 ± 1.1	-46.6 ± 4	**-55.2 ± 18.2**	eu	-74.9 ± 1.9	-67.3 ± 3.3	**-73.1 ± 18.2**
1,1-dmu	-51.6 ± 1.7	-48.4 ± 4.3	**-49.1 ± 15.9**	1,1-dmu	-60.2 ± 1.7	-62.5 ± 3.6	**-68 ± 15.9**
1,3-dmu	-52.6 ± 1.6	-43.5 ± 5.5	**-48.2 ± 13.2**	1,3-dmu	-66.7 ± 1.9	-74.5 ± 4.6	**-71 ± 13.2**
1,1-deu	-51.4 ± 1.7	-46.6 ± 6.1	**-53.3 ± 15.3**	1,1-deu	-77.5 ± 2.3	-81.9 ± 5	**-80.2 ± 15.3**
1,3-deu	-48.9 ± 1.3	-42.4 ± 6.1	**-54.4 ± 13.1**	1,3-deu	-81.6 ± 2.5	-87.1 ± 5	**-85.9 ± 13.1**
mad	**-54.1 ± 2.4**	**-65.3 ± 8.2**	**-63.9 ± 27.3**	**mad**	**-64.5 ± 3**	**-52.9 ± 7.2**	**-71.8 ± 27.3**
ppa	**-76.8 ± 2.6**	**-82.6 ± 11.4**	**-46.2 ± 15.7**	**ppa**	**-102.3 ± 1.9**	**-96.3 ± 9.7**	**-62.8 ± 15.7**
aama	**-53 ± 3.1**	**-60.6 ± 23**	**-54.5 ± 18.9**	**aama**	**-80.3 ± 3.2**	**-71.4 ± 17.9**	**-82.2 ± 18.9**
mfad	**-64 ± 2.2**	**-63.7 ± 23**	**-63.1 ± 17.9**	**fad**	**-64 ± 2.2**	**-63.1 ± 17.9**	**-62.5 ± 23**
	aama (Solute 3)				**aama** (Solute 3)		
Solute 2	Observed μ_{23}[a]	Predicted μ_{32} or μ_{23} from one-way alpha values of Solute 2[b]	Solute 3[c]	Solute 2	Observed μ_{23}[a]	Predicted μ_{32} or μ_{23} from one-way alpha values of Solute 2[b]	Solute 3[c]
urea	-52.4 ± 5.9	-54 ± 2.5	**-47.4 ± 23.4**	1,3-deu	-80.3 ± 2.9	-79.9 ± 7	**-88.8 ± 13.1**
mu	-61.3 ± 3.3	-61.9 ± 5.9	**-58.8 ± 18.3**	**mad**	**-8.6 ± 2.2**	**-10.4 ± 7.3**	**-12.9 ± 27.3**
eu	-70.6 ± 1.7	-76.9 ± 4.4	**-68.1 ± 18.2**	**ppa**	**-43.2 ± 3.9**	**-43.2 ± 9.3**	**-39.8 ± 15.7**
1,1-dmu	-39.8 ± 2.3	-41.3 ± 4.7	**-43.1 ± 15.9**	**fad**	**-53 ± 3.1**	**-54.5 ± 18.9**	**-60.6 ± 23**
1,3-dmu	-60.9 ± 2.6	-63.2 ± 6.3	**-70.2 ± 13.2**	**mfad**	**-80.3 ± 3.2**	**-78.4 ± 18.9**	**-71.4 ± 17.9**
1,1-deu	-78.3 ± 3.1	-82.1 ± 7	**-64.6 ± 15.3**				

Footnotes to Table S3 are on the next page

[a] New μ_{23} values (Table S1) and predictions from new one-way alpha values (Table S2) are in bold font. Amide-amide interactions determined by VPO at 23°C (Table S1; ref. (1)). Amide-aromatic interactions determined by solubility assay at 25°C (ref. (2)). Uncertainties are from fitting μ_{23} data to Eq. S17 using Igor Pro (WaveMetrics, Lake Oswego, OR, USA).

[b] Calculated from Eq. 1 using one-way alpha values of solute 2 in Table S5 and this paper (Table S2) and ASA information (1). One-way alpha values in Table S5 for urea and alkyl ureas differ slightly from those in ref. (1) because they are calculated using the combined set of amide and aromatic μ_{23} values. Propagated uncertainties were calculated as in ref. (2).

[c] Calculated from Eq. 1 using one-way alpha values of solute 3 in Table S2 and ASA information (1). Propagated uncertainties were calculated as in reference (2).

Abbreviations: mu: methylurea; eu: ethylurea; dmu: dimethylurea; deu: diethylurea; mad: malonamide; ppa: propionamide; aama: N-acetylalanine N-methylamide; fad: formamide; mfad: N-methylformamide.

Table S4. Comparison of Observed μ_{23} Values (cal mol^{-1} molal^{-1}) for Amide Interactions with Predictions from Two-Way Alpha Values (Table 1)

Solute 2	Solute 3	Observed μ_{23}[a]	Predicted μ_{23}[b]	Solute 2	Solute 3	Observed μ_{23}[a]	Predicted μ_{23}[b]
urea	mu	-37.8 ± 1.9	-37.5 ± 20.1	mu	eu	-59.4 ± 2.3	-62.4 ± 15.3
urea	eu	-43.8 ± 2.3	-39 ± 19.8	mu	1,1-dmu	-46.8 ± 2.1	-57.8 ± 13.7
urea	1,1-dmu	-35.7 ± 2.1	-38.5 ± 17.5	mu	1,3-dmu	-59.1 ± 2.8	-68.2 ± 11.4
urea	1,3-dmu	-30.2 ± 1.1	-32.9 ± 14.3	mu	1,1-deu	-78.3 ± 1.9	-73.9 ± 13.6
urea	1,1-deu	-39.7 ± 1.2	-38.3 ± 17	mu	1,3-deu	-87.8 ± 3.3	-87.3 ± 11.9
urea	1,3-deu	-40.3 ± 2.5	-35.9 ± 14.7	mu	mad	-43.8 ± 1.4	-41.3 ± 23.8
urea	mad	-54.9 ± 2.5	-56.3 ± 31.5	mu	ppa	-62.3 ± 1.7	-51.2 ± 13.5
urea	ppa	-36.8 ± 2	-35.8 ± 17.4	mu	nma	-49.7 ± 1.8	-50.2 ± 11.4
urea	nma	-36.3 ± 1.7	-34.6 ± 14.3	mu	aama	-61.3 ± 3.3	-61.7 ± 18.8
urea	aama	-52.4 ± 5.9	-56.5 ± 23.7	mu	fad	-48.8 ± 3	-54.2 ± 18.8
urea	fad	-62.9 ± 2.2	-57.6 ± 24.8	mu	nma	-53.7 ± 1.7	-69.3 ± 14.9
urea	mfad	-56.1 ± 2.9	-53 ± 19.1	mu	naphthalene	-361 ± 15	-370.1 ± 51.3
urea	SDS	-22.4 ± 1.2	-20.9 ± 14	mu	anthracene	-475 ± 43	-452.8 ± 62.8
urea	butane	-11.8 ± 1.2	-11.9 ± 7.9	eu	1,1-dmu	-69.2 ± 2.7	-69.2 ± 13.8
urea	CycloGlyGly	-47.7 ± 2.4	-72.5 ± 34.2	eu	1,3-dmu	-83.1 ± 2.9	-85.7 ± 11.5
urea	CycloAlaGly	-64.7 ± 5.9	-66.2 ± 30.3	eu	1,1-deu	-96.2 ± 2.7	-92.1 ± 13.7
urea	CycloAlaAla	-64.7 ± 5.9	-66.9 ± 30.5	eu	1,3-deu	-102.7 ± 1.9	-111.8 ± 12.1
urea	CycloGlyLeu	-70.6 ± 5.9	-71.9 ± 31.7	eu	mad	-48.5 ± 2.4	-40 ± 23.5
urea	CycloValVal	-70.6 ± 5.9	-64.7 ± 27	eu	ppa	-67.6 ± 1.7	-60.8 ± 13.6
urea	AG2EE-AGEE	-34.1 ± 2.9	-23.1 ± 10.1	eu	nma	-72.2 ± 1.5	-59.9 ± 11.6
urea	AG3EE-AGEE	-57.7 ± 4.7	-48.2 ± 21.3	eu	aama	-70.6 ± 1.7	-69.3 ± 19.1
urea	AG4EE-AGEE	-87.1 ± 5.9	-72.2 ± 32	eu	fad	-39.2 ± 1.1	-57.5 ± 18.6
urea	ALEE-AGEE	6.5 ± 0.6	-1.3 ± 5.1	eu	mfad	-74.9 ± 1.9	-80.5 ± 14.9
urea	AAEE-AGEE	14.7 ± 1.2	4.5 ± 4.2	eu	naphthalene	-390 ± 15	-475 ± 51
urea	naphthalene	-166 ± 5.9	-139 ± 67	eu	anthracene	-588 ± 40	-581 ± 63
urea	anthracene	-196 ± 5.9	-171 ± 82				

1,1-dmu	1,3-dmu	-61.6 ± 1.9	-77 ± 10.6	1,1-deu	mfad	-77.5 ± 2.3	-91.7 ± 13.7
1,1-dmu	1,1-deu	-77.7 ± 2.7	-79.4 ± 12.7	1,1-deu	naphthalene	-748.6 ± 13.9	-709 ± 48.5
1,1-dmu	1,3-deu	-79.8 ± 2.8	-100 ± 11.4	1,1-deu	anthracene	-913.5 ± 56.5	-867.4 ± 59.3
1,1-dmu	mad	-30.6 ± 1.8	-27.8 ± 21.1	1,3-deu	mad	-21.7 ± 1.2	-23.8 ± 17.9
1,1-dmu	ppa	-53.5 ± 2.3	-51 ± 12.4	1,3-deu	ppa	-72.7 ± 2.5	-85.9 ± 11
1,1-dmu	nma	-42.2 ± 1.7	-42 ± 10.9	1,3-deu	nma	-87.1 ± 2.2	-85.3 ± 10.2
1,1-dmu	aama	-39.8 ± 2.3	-35.6 ± 18	1,3-deu	aama	-80.3 ± 2.9	-82.2 ± 16.7
1,1-dmu	fad	-51.6 ± 1.7	-53.7 ± 16.8	1,3-deu	fad	-48.9 ± 1.3	-57.3 ± 14.5
1,1-dmu	mfad	-60.2 ± 1.7	-72.5 ± 13.8	1,3-deu	mfad	-81.6 ± 2.5	-108 ± 13
1,1-dmu	naphthalene	-554 ± 33	-555 ± 48	1,3-deu	naphthalene	-743 ± 16	-810 ± 45
1,1-dmu	anthracene	-697 ± 33	-679 ± 59	1,3-deu	anthracene	-1071 ± 75	-991 ± 55
1,3-dmu	1,1-deu	-114 ± 3.6	-110 ± 11	**mad**	**ppa**	-37.3 ± 2.2	-28.2 ± 20.8
1,3-dmu	1,3-deu	-125 ± 3.5	-139 ± 9.8	**mad**	**aama**	-8.6 ± 2.2	-7.4 ± 28.8
1,3-dmu	mad	-27.3 ± 2.4	-26.4 ± 17.3	**mad**	**fad**	-54.1 ± 2.4	-70.6 ± 29.4
1,3-dmu	ppa	-68.5 ± 1.9	-66.7 ± 10.3	**mad**	**mfad**	-64.5 ± 3	-55.5 ± 23
1,3-dmu	nma	-66.2 ± 2.8	-65.7 ± 9.3	**ppa**	**aama**	-43.2 ± 3.9	-34.4 ± 17.3
1,3-dmu	aama	-60.9 ± 2.6	-66.7 ± 15.2	**ppa**	**fad**	-76.8 ± 2.6	-49.7 ± 16.4
1,3-dmu	fad	-52.6 ± 1.6	-50.7 ± 13.8	**ppa**	**mfad**	-102 ± 1.9	-64.8 ± 13.3
1,3-dmu	mfad	-66.7 ± 1.9	-85.5 ± 11.6	**aama**	**fad**	-53 ± 3.1	-67.2 ± 23.3
1,3-dmu	naphthalene	-569 ± 9.4	-601 ± 40.4	**aama**	**mfad**	-80.3 ± 3.2	-71.5 ± 20
1,3-dmu	anthracene	-639 ± 24	-735 ± 49.4	**fad**	**mfad**	-64 ± 2.2	-75.3 ± 20.8
1,1-deu	1,3-deu	-146 ± 4.5	-146 ± 11.7	eeu	naphthalene	-471 ± 24	-498 ± 57
1,1-deu	mad	-34.6 ± 1.3	-25.3 ± 20.6	tmu	naphthalene	-918 ± 25	-872 ± 45
1,1-deu	ppa	-78.2 ± 1.5	-69.1 ± 12.3	dfad	naphthalene	-678 ± 20	-695 ± 55
1,1-deu	nma	-89 ± 1.9	-63.2 ± 11.1	ndma	naphthalene	-761 ± 38	-763 ± 46
1,1-deu	aama	-78.3 ± 3.1	-56.5 ± 18.3	mfad	naphthalene	-538 ± 25	-513 ± 58
1,1-deu	fad	-51.4 ± 1.7	-56.7 ± 16.3	acet	naphthalene	-380 ± 12	-405 ± 53

Footnotes to Table S4 are on next page

[a]Amide-amide interactions determined by VPO at 23°C (Table S1; (1)). Amide-aromatic interactions determined by solubility assay at 25°C (2). Uncertainties in fitting VPO and solubility data were determined as described previously (1, 2).

[b]Calculated from Eq.3 using two-way alpha values from Table 1 and ASA information (1). Uncertainties are calculated from propagated uncertainties in two-way alpha values (Table 1).

Abbreviations: **aa**: acetamide; **dfad**: N, N-dimethylformamide; **tmu**: tetramethyl urea; **eeu**: ethyleneurea; **ndma**: N,N-dimethyl acetamide. Other solute abbreviations are listed in Table S2

Table S5. Comparison of One-way Alpha Values (cal mol^{-1} molal^{-1} Å^{-2}) Calculated from μ_{23} Values with Predictions of One-way Alpha Values from Two-way Alpha Values

Solute [a]	One-way Alpha Value (cal mol^{-1} m^{-1} Å^{-2})							
	Amide sp^{2}O [b]		Amide sp^{2}N [b]		Amide sp^{3}C [b]		Aromatic/Amide sp^{2}C [b]	
	Calculated from μ_{23} [c]	Predicted [d]	Calculated from μ_{23} [c]	Predicted [d]	Calculated from μ_{23} [c]	Predicted [d]	Calculated from μ_{23} [c]	Predicted [d]
urea	-0.53 ± 0.03	-0.64 ± 0.33	-0.1 ± 0.02	-0.06 ± 0.17	-0.06 ± 0	-0.05 ± 0.05	-0.59 ± 0.01	-0.51 ± 0.37
mu	0.78 ± 0.08	0.61 ± 0.25	-0.51 ± 0.03	-0.44 ± 0.14	-0.34 ± 0.01	-0.33 ± 0.04	-1.36 ± 0.08	-1.36 ± 0.32
eu	1.08 ± 0.07	1.06 ± 0.25	-0.62 ± 0.02	-0.59 ± 0.15	-0.46 ± 0.01	-0.45 ± 0.05	-1.6 ± 0.07	-1.74 ± 0.33
1,1-dmu	1.45 ± 0.07	1.63 ± 0.23	-0.69 ± 0.03	-0.78 ± 0.13	-0.39 ± 0.01	-0.43 ± 0.04	-2.03 ± 0.07	-2.03 ± 0.3
1,3-dmu	1.74 ± 0.06	1.86 ± 0.19	-0.78 ± 0.03	-0.82 ± 0.11	-0.57 ± 0.01	-0.61 ± 0.04	-1.95 ± 0.04	-2.2 ± 0.26
1,1-deu	2.31 ± 0.09	2.31 ± 0.23	-0.98 ± 0.03	-1 ± 0.13	-0.7 ± 0.01	-0.66 ± 0.04	-2.69 ± 0.1	-2.6 ± 0.31
1,3-deu	2.89 ± 0.12	2.77 ± 0.21	-1.14 ± 0.04	-1.13 ± 0.12	-0.83 ± 0.02	-0.85 ± 0.04	-2.97 ± 0.13	-2.97 ± 0.28
Solute [a]	Amide sp^{2}O [e]		Amide sp^{2}N [e]		Amide sp^{3}C [e]		Amide sp^{2}C [e]	
	Calculated from μ_{23} [c]	Predicted [d]	Calculated from μ_{23} [c]	Predicted [d]	Calculated from μ_{23} [c]	Predicted [d]	Calculated from μ_{23} [c]	Predicted [d]
fad	-0.13 ± 0.08	-0.4 ± 0.32	-0.33 ± 0.03	-0.24 ± 0.18	-0.14 ± 0.01	-0.13 ± 0.06	-0.92 ± 0.08	-1.03 ± 0.42
mfad	-0.02 ± 0.09	0.86 ± 0.27	-0.41 ± 0.03	-0.62 ± 0.15	-0.27 ± 0.01	-0.4 ± 0.05	-0.81 ± 0.09	-1.88 ± 0.36
mad	0.37 ± 0.07	0.26 ± 0.4	-0.5 ± 0.02	-0.46 ± 0.21	-0.07 ± 0.01	-0.03 ± 0.06	-1.23 ± 0.07	-1.28 ± 0.45
ppa	1.04 ± 0.09	1.29 ± 0.22	-0.56 ± 0.04	-0.65 ± 0.13	-0.33 ± 0.01	-0.36 ± 0.04	-2.39 ± 0.1	-1.73 ± 0.29
aama	2.88 ± 0.21	3.63 ± 0.35	-1.26 ± 0.09	-1.57 ± 0.17	-0.45 ± 0.01	-0.45 ± 0.06	-2.97 ± 0.14	-3.53 ± 0.4

[a] See Table S3-4 for abbreviations.

[b] Calculated from μ_{23} values for amide-amide and amide-aromatic interactions assuming no distinction between aromatic sp^{2}C and amide sp^2 C (see Table S6).

[c] Calculated from Eq. 1.

[d] Predicted from two-way alpha values and ASA information using Eq. 2 .

[e] Calculated from μ_{23} values for amide-amide interactions only.

Table S6. Comparison of Two-way Alpha Values for Atom-atom Interactions Using Different Treatments of Amide and Aromatic sp^2C

atom i	atom j	Two-way Alpha Values (millical mol^{-1} $molal^{-1}$ $Å^{-4}$) from Analysis of			
		105 μ_{23} values[a]	105 μ_{23} values[b] (1 global weight)	64 amide-amide μ_{23} values[c]	20 amide-aromatic μ_{23} values[d]
sp^2O	sp^2C	-13.9 ± 3.8	-14 ± 3.8	ND	-16.3 ± 2.9
sp^2O	sp^2N	-10.8 ± 1.7	-8.1 ± 1.7	-12.4 ± 1.6	ND
sp^2C	sp^3C	-11.1 ± 3.3	-13.5 ± 3.6	ND	-10.6 ± 3
sp^2C	sp^2C	-10.3 ± 0.5	-10.3 ± 0.5	ND	-9.9 ± 0.3
sp^3C	sp^3C	-3.9 ± 0.1	-3.6 ± 0.1	-3.7 ± 0.1	ND
sp^3C	sp^2N	-3.8 ± 0.2	-3.4 ± 0.2	-3.5 ± 0.2	ND
sp^2C	sp^2N	1.8 ± 1.3	1.8 ± 1.3	ND	3.1 ± 1.1
sp^2N	sp^2N	3.4 ± 0.6	2.2 ± 0.6	3.9 ± 0.7	ND
sp^3C	sp^2O	10.8 ± 0.5	9.5 ± 0.5	8.9 ± 0.5	ND
sp^2O	sp^2O	18.1 ± 5.1	10.5 ± 4.9	18.7 ± 4.2	ND
weight of amide sp^2C relative to aromatic sp^2C		1	0.82		
ssresid[e]		48533	46898		

[a] Two-way alpha values from Table 1, from analysis of 105 μ_{23} values for amide-amide and amide-aromatic interactions by Eq. 3, combining amide and aromatic sp^2C.

[b] Two-way alpha values from analysis of 105 μ_{23} values for amide-amide and amide-aromatic interactions by Eq. S15, in which interactions of amide sp^2C are assumed to differ from aromatic sp^2C by a common global weighing factor.

[c] Two-way alpha values from analysis of 64 μ_{23} values for amide-amide interactions from Tables S3-4 by Eq. 3, including malonamide, propionamide and aama (Table S1) but excluding compounds with large sp^2C ASA (fad, mfad, aromatics) and combining small amounts of amide sp^2C ASA (less than 6 $Å^2$) with aliphatic sp^3C ASA.

[d] Two-way amide-aromatic alpha values obtained directly from naphthalene one-way alpha values using naphthalene ASA (273 $Å^2$) (1).

[e] Ssresid: Sum of squares of residuals for predicted - observed differences in the set of μ_{23} values $\sum(\mu_{23}^{predicted} - \mu_{23}^{observed})^2$

Table S7. Comparison of Two-way Alpha Values Obtained by Fitting with Estimates from One-way Alpha Values for Amide Pairs Differing Primarily in ASA of One Atom Type

Amide Pair[a]	%ΔASA[b]	Two-way Alpha Value (cal mol^{-1} molal^{-1} Å^{-4})							
		sp^2O		sp^2N		sp^2C		sp^3C	
		Estimate (Eq. S1)	Fitted (Table 1)	Estimate (Eq. S1)	Fitted (Table 1)	Estimate (Eq. S1)	Fitted (Table 1)	Estimate (Eq. S1)	Fitted (Table 1)
eu–mu	(86% sp^3C)[c]	0.008	0.011	-0.003	-0.0038	-0.006	-0.010	-0.003	-0.0039
1,1-deu – 1,1-dmu	(90% sp^3C)[d]	0.010	0.011	-0.004	-0.0038	-0.006	-0.010	-0.004	-0.0039
1,3-deu – ppa	(87% sp^2N)[e]	-0.012	-0.011	0.001	0.0034	(0.058)	0.0018	-0.004	-0.0038
aama – 1,3-deu	(63% sp^2O)[f]	0.016	0.018	-0.009	-0.011	-0.020	-0.014	0.009	0.011

[a] Abbreviations for amide compounds as in Table S3-4.

[b] ΔASA calculated as the sum of magnitudes of ASA differences for the different atom types.

[c] ΔASA = 42 Å^2 (36 Å^2 sp^3C, -6 Å^2 sp^2N)

[d] ΔASA = 71 Å^2 (64 Å^2 sp^3C, -4 Å^2 sp^2N, -3 Å^2 sp^2O)

[e] ΔASA = 23 Å^2 (20 Å^2 sp^2N, 2 Å^2 sp^3C, 1 Å^2 sp^2O)

[f] ΔASA = 54 Å^2 (34 Å^2 sp^2O, -13 Å^2 sp^2N; 7 Å^2 sp^3C)

Table S8. PVP Oligomer ASA Values in Å^2 from Cactus Models; Comparison with PubChem

Number of oligomer residues N_3	Source	Amide sp^2O	Amide sp^2N	Aliphatic sp^3C	Amide sp^2C	Total
1 (NEP) [a]	Cactus	36.9	0.35	252	2.6	292
2 [a]		54.5	0.35	381	3.0	439
3 [a]		81.5	0.35	479	3.4	564
4 [a]		88.9	0.35	578	3.8	671
5 [a]		97.5	0.35	674	3.4	775
Interior residue [b]	Cactus (linear fitting)	15.6	0	104	0.2	120
End residues [b]		56.3	0.35	369	3.0	428
1 (NEP) [c]	PubChem	39.0	0.52	250	2.8	292
2 [c]		54.6	0.35	381	3.0	439

[a] Molecular models created as described in SI text. Chemical formula of interior PVP residue is C_6H_9NO, end residue $C_6H_{10}NO$

[b] Determined with Eq. S16 as described in SI

[c] Molecular models created with PubChem

Table S9. PVP One-Way End and Interior Residue Alpha Values; Comparison with NEP and PEG

Protein Atom Type	NEP (Monomer) $(\alpha_{NEP,i})$ [a]	PVP Interior Residue $(\alpha_{I,i})$ [a]	PEG Interior Residue $(\alpha_{I,i})$ [a]	PVP End Residues $(\alpha_{E,i})$ [a]	PEG End Residues $(\alpha_{2E,i})$ [a]
Aliphatic sp^3C	0.61 ± 0.03	-0.24 ± 0.01	-0.12 ± 0.01	-0.86 ± 0.04	-0.003 ± 0.02
Amide sp^2O	3.4 ± 0.2	1.4 ± 0.1	0.75 ± 0.04	5.0 ± 0.3	0.70 ± 0.05
Amide sp^2N	-1.4 ± 0.1	-0.50 ± 0.03	-0.42 ± 0.02	-2.0 ± 0.1	-0.42 ± 0.03
Amide sp^2C	-3.1 ± 0.2	-1.3 ± 0.07	-0.77 ± 0.01	-4.6 ± 0.3	-0.36 ± 0.09

[a] $cal^{-1}\ mol^{-1}\ molal^{-1}\ Å^{-2}$

Table S10. Predicted Chemical Contributions to m-Values Quantifying Effects of NEP (PVP Model Monomer) and PVP Polymer Residues on Protein and α-Helix Unfolding

	Globular protein unfolding	α-helix unfolding
NEP Monomer	-200 ± 40[a]	1600 ± 100[a]
PVP Interior residue	-70 ± 20[b]	690 ± 60[b]
PVP End residue	-270 ± 70[b]	2400 ± 200[b]

[a] cal mol^{-1} $molal^{-1}$ for interaction with 1000 $Å^2$ of protein ΔASA.

[b] cal mol^{-1} $molal^{-1}$ $residue^{-1}$ for 1000 $Å^2$ of protein ΔASA.

Table S11. Test of Dataset Size by Excluding μ_{23} Values for Selected Polar (urea, mad) and Nonpolar (1,3-deu, aama) Amides[a]

Amide Atom		Two-way Alpha Values (millical mol^{-1} $molal^{-1}$ $Å^{-4}$)				
i	j	Entire set of 105 μ_{23}	Exclude urea (79 μ_{23})	Exclude 1,3-deu (91 μ_{23})	Exclude aama (94 μ_{23})	Exclude 1,3-deu and mad (81 μ_{23})
sp^2O	sp^2C	-13.9 ± 3.8	-13.4 ± 3.9	-15.3 ± 5.4	-14 ± 3.9	-15.3 ± 3.3
sp^2O	sp^2N	-10.8 ± 1.7	-10.7 ± 2.1	-12.5 ± 6.5	-9.8 ± 2.3	-11.4 ± 1.8
sp^2C	sp^3C	-10.3 ± 0.5	-10.5 ± 0.5	-10.1 ± 0.6	-10.3 ± 0.5	-10.1 ± 0.4
sp^2C	sp^2C	-11.1 ± 3.3	-11.2 ± 3.4	-10.3 ± 5	-11.1 ± 3.4	-10.3 ± 3.1
sp^3C	sp^3C	-3.9 ± 0.1	-3.9 ± 0.1	-4 ± 0.4	-3.9 ± 0.1	-4.1 ± 0.1
sp^3C	sp^2N	-3.8 ± 0.2	-3.7 ± 0.2	-3.6 ± 0.8	-4.1 ± 0.2	-3.8 ± 0.2
sp^2C	sp^2N	1.8 ± 1.3	2.5 ± 1.4	2.2 ± 1.8	1.8 ± 1.3	2.2 ± 1.2
sp^2N	sp^2N	3.4 ± 0.6	3.2 ± 0.9	4 ± 2.8	3 ± 0.9	3.8 ± 0.6
sp^3C	sp^2O	10.8 ± 0.5	11 ± 0.5	10.3 ± 2.2	11.6 ± 0.7	11 ± 0.6
sp^2O	sp^2O	18.1 ± 5.1	16.9 ± 5.4	23.2 ± 17.2	15.5 ± 6.7	19.8 ± 5.3

[a] Abbreviations as in Tables S3-4.

References

1. X. Cheng *et al.*, Experimental Atom-by-Atom Dissection of Amide-Amide and Amide-Hydrocarbon Interactions in H2O. *J. Am. Chem. Soc.* **139**, 9885-9894 (2017).

2. D. B. Knowles *et al.*, Chemical Interactions of Polyethylene Glycols (PEGs) and Glycerol with Protein Functional Groups: Applications to Effects of PEG and Glycerol on Protein Processes. *Biochemistry* **54**, 3528-3542 (2015).

3. I. A. Shkel, D. B. Knowles, M. T. Record, Separating chemical and excluded volume interactions of polyethylene glycols with native proteins: Comparison with PEG effects on DNA helix formation. *Biopolymers* **103**, 517-527 (2015).

4. D. B. Knowles, A. S. LaCroix, N. F. Deines, I. Shkel, M. T. Record, Separation of preferential interaction and excluded volume effects on DNA duplex and hairpin stability. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12699-12704 (2011).

5. A. C. Miklos, C. G. Li, G. J. Pielak, "Using NMR-detected backbone amide H-1 exchange to assess macromolecular crowding effects on globular-protein stability" in Methods in Enzymology, Vol 466: Biothermodynamics, Pt B*,* M. L. Johnson, J. M. Holt, G. K. Ackers, Eds. (Elsevier Academic Press Inc, San Diego, 2009), vol. 466, pp. 1-18.

6. L. M. Charlton *et al.*, Residue-level interrogation of macromolecular crowding effects on protein stability. *Journal of the American Chemical Society* **130**, 6826-6830 (2008).

7. A. C. Miklos, C. G. Li, N. G. Sharaf, G. J. Pielak, Volume Exclusion and Soft Interaction Effects on Protein Stability under Crowded Conditions. *Biochemistry* **49**, 6984-6991 (2010).

8. E. J. Guinn, L. M. Pegram, M. W. Capp, M. N. Pollock, M. T. Record, Quantifying why urea is a protein denaturant, whereas glycine betaine is a protein stabilizer. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16932-16937 (2011).

9. E. J. Guinn *et al.*, Quantifying Functional Group Interactions That Determine Urea Effects on Nucleic Acid Helix Formation. *J. Am. Chem. Soc.* **135**, 5828-5838 (2013).

10. R. J. Hefford, Polymer Mixing in Aqueous-Solution. *Polymer* **25**, 979-984 (1984).

11. K. A. Dill, S. Bromberg, D. Stigter, *Molecular driving forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience 2* (Garland Science, 2012).

12. J.-M. Choi, A. S. Holehouse, R. V. Pappu, Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annual Review of Biophysics* **49**, 107-133 (2020).

13. M. Feric *et al.*, Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686-1697 (2016).

14. H. Hofmann *et al.*, Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16155-16160 (2012).

15. A. Soranno *et al.*, Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4874-4879 (2014).

16. J. A. Riback *et al.*, Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* **168**, 1028-1014 (2017).

17. Y. Shin, C. P. Brangwynne, Liquid phase condensation in cell physiology and disease. *Science* **357**, 7 (2017).

18. T. M. Franzmann *et al.*, Phase separation of a yeast prion protein promotes cellular fitness. *Science* **359**, 47 (2018).

19. J. Sadowski, J. Gasteiger, G. Klebe, Comparison of Automatic 3-Dimensional Model Builders Using 639 X-Ray Structures. *J Chem Inf Comp Sci* **34**, 1000-1008 (1994).

20. O. V. Tsodikov, M. T. Record, Jr., Y. V. Sergeev, Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *Journal of computational chemistry* **23**, 600-609 (2002).

21. J. R. Livingstone, R. S. Spolar, M. T. Record, Jr., Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. *Biochemistry* **30**, 4237-4244 (1991).

22. E. E. Bolton *et al.*, PubChem3D: a new resource for scientists. *J Cheminformatics* **3** (2011).

23. E. L. Ulrich *et al.*, BioMagResBank. *Nucleic Acids Res* **36**, D402-D408 (2008).

24. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).

25. R. Fraczkiewicz, W. Braun, Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *Journal of computational chemistry* **19**, 319-333 (1998).

26. X. Cheng *et al.*, Quantifying Interactions of Nucleobase Atoms with Model Compounds for the Peptide Backbone and Glutamine and Asparagine Side Chains in Water. *Biochemistry* **57**, 2227-2237 (2018).